

# Analysis III

Winter 2017/2018



Totales Differential, Taylor-Polynom, Mittelwertsatz

Buch Kapitel 5.10-5.11

# Erinnerung: Ableitungsmatrix und Lineare Approximation

**Definition (Ableitungsmatrix):**

Sei  $f: D \rightarrow \mathbb{R}^m$ ,  $D \subset \mathbb{R}^n$  eine Abbildung mit

$$f(\mathbf{x}) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in D,$$

die in  $\mathbf{x}_0$  partiell diff'bar ist. Dann existieren die Ableitungen

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0), \quad i = 1, \dots, m, \quad j = 1, \dots, n,$$

und die **Ableitungsmatrix** ist definiert als

$$f'(\mathbf{x}_0) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_m}{\partial x_{n-1}}(\mathbf{x}_0) & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}.$$

Andere Bezeichnungen für diese Matrix: **Ableitung** von  $f$  oder **Jacobi-Matrix**.

**Bemerkungen (Voraussetzungen):**

- Sei  $f: D \rightarrow \mathbb{R}^m$ ,  $D \subset \mathbb{R}^n$ ,  $\mathbf{x}_0 \in D$ . Dann gilt:
  - $f'(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}$
- Die Ableitungsmatrix ist **linear** (in  $\mathbf{x}$ ):
  - $f'(\mathbf{x}_0) = f'(\mathbf{x}_0) \cdot \mathbf{1} = f'(\mathbf{x}_0) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) + \dots + \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) + \dots + \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}$
- Die Ableitungsmatrix ist **symmetrisch** (in  $\mathbf{x}$ ):
  - $f'(\mathbf{x}_0) = f'(\mathbf{x}_0) \cdot \mathbf{1} = f'(\mathbf{x}_0) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) + \dots + \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) + \dots + \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}$
- Die Ableitungsmatrix ist **symmetrisch** (in  $\mathbf{x}$ ):
  - $f'(\mathbf{x}_0) = f'(\mathbf{x}_0) \cdot \mathbf{1} = f'(\mathbf{x}_0) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) + \dots + \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) + \dots + \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}$

**Definition (Hessematrix):**

Sei  $f: D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  eine skalarwertige in  $\mathbf{x} \in D$  zweifach partiell differenzierbare Funktion mit Ableitungen

$$\frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right) = \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad i, j = 1, \dots, n.$$

Die **Hesse-Matrix** der Funktion  $f$  ist definiert durch

$$H_f(\mathbf{x}) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_{n-1}}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{pmatrix}.$$

**Vorbemerkung:**

In komplizierten nichtlinearen funktionalen Zusammenhängen interessiert eine (näherungsweise) lineare Darstellung.  
Frage: Kann man die finden?

**Idee:**

Ist  $f: D \rightarrow \mathbb{R}^m$ ,  $D \subset \mathbb{R}^n$  differenzierbare Abbildung, so lässt sie sich in der Nähe von  $\mathbf{x}_0 \in D$  durch  $g: D \rightarrow \mathbb{R}^m$  darstellen durch

$$g(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).$$

Nach der Definition der Diff'barkeit begeht man dabei einen "kleinen" Fehler ( $\|k(\mathbf{x})\|$ ).

**Definition:**

Die eben definierte Abbildung  $g(\mathbf{x})$  nennt man **Tangentialabbildung** oder **lineare Approximation** von  $f$  in  $\mathbf{x}_0$ .

**Definition** (Ableitungsmatrix):

Sei  $\mathbf{f} : D \rightarrow \mathbb{R}^m$ ,  $D \subset \mathbb{R}^n$  eine Abbildung mit

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in D,$$

die in  $\mathbf{x}_0$  partiell diff'bar ist. Dann existieren die Ableitungen

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0), \quad i = 1, \dots, m, \quad j = 1, \dots, n,$$

und die **Ableitungsmatrix** ist definiert als

$$\mathbf{f}'(\mathbf{x}_0) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_m}{\partial x_{n-1}}(\mathbf{x}_0) & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}.$$

Andere Bezeichnungen für diese Matrix: *Ableitung von  $\mathbf{f}$*  oder *Jacobi-Matrix*.

**Definition** (Hessematrix):

Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  eine skalarwertige in  $\mathbf{x} \in D$  zweifach partiell differenzierbare Funktion mit Ableitungen

$$\frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right) = \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad i, j = 1, \dots, n.$$

Die **Hesse-Matrix** der Funktion  $f$  ist definiert durch

$$H_f(\mathbf{x}) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_{n-1}}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}) \end{pmatrix}.$$

### Vorbemerkung:

In komplizierten nichtlinearen funktionalen Zusammenhängen interessiert eine (näherungsweise) Lineare Darstellung.

Frage: Kann man die finden?

### Idee:

Ist  $\mathbf{f} : D \rightarrow \mathbb{R}^m$ ,  $D \subset \mathbb{R}^n$  differenzierbare Abbildung, so lässt sie sich in der Nähe von  $\mathbf{x}_0 \in D$  durch  $\mathbf{g} : D \rightarrow \mathbb{R}^m$  darstellen durch

$$\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).$$

Nach der Definition der Diff'barkeit begeht man dabei einen "kleinen" Fehler ( $|\mathbf{k}(\mathbf{x})|$ ).

### Definition:

Die eben definierte Abbildung  $\mathbf{g}(\mathbf{x})$  nennt man **Tangentialabbildung** oder **lineare Approximation** von  $\mathbf{f}$  in  $\mathbf{x}_0$ .



### Bemerkungen (Tangentialebene):

- Betrachte  $f : D \rightarrow \mathbb{R}$  mit  $D \subset \mathbb{R}^2$ , in  $\mathbf{x}_0 = (x_0, y_0)^\top$  diff'bar. Dann gilt:

$$\begin{aligned} f(x, y) &= f(x_0, y_0) + [f_x(x_0, y_0), f_y(x_0, y_0)] \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + k(x, y) \\ &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + k(x, y). \end{aligned}$$

- Dementsprechend lautet die Tangentenabbildung

$$g(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

- Der Graph von  $g$  ist eine Ebene, die **Tangentialebene**.
- die Tangentialebene berührt in  $P = (x_0, y_0, g(x_0, y_0))$  den Graphen der Funktion  $f$ , also

$$(x_0, y_0, g(x_0, y_0)) = (x_0, y_0, f(x_0, y_0))$$

In der Umgebung von  $(x_0, y_0)$  nähert sich  $g$  der Funktion  $f$  an.

- Die beiden Tangenten

$$\begin{aligned} z &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) \quad \text{und} \\ z &= f(x_0, y_0) + f_y(x_0, y_0)(y - y_0) \end{aligned}$$

an die Funktionen  $f^*(x) := f(x, y_0)$  bzw.  $f^{**}(y) = f(x_0, y)$  liegen in der Tangentialebene.

- Wegen  $k(x, y)$  klein, stellt die Ebene eine gute Näherung der Funktion  $f$  in der Umgebung von  $(x_0, y_0)$  dar.

# Totales Differential

## Idee:

Verwende die allgemeine Definition von  $f'(\mathbf{x})$ :

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + k(\mathbf{x}) \\ \Rightarrow f(\mathbf{x}) - f(\mathbf{x}_0) = \Delta z &= f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + k(\mathbf{x}). \end{aligned}$$

**Bemerkung:** Da  $|k(\mathbf{x})|$  sehr klein ist für  $|\mathbf{x} - \mathbf{x}_0|$  klein, ist  $\Delta z \approx f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$  dann eine gute Näherung.

## Definition: (totales Differential)

Verwende folgende Notation:

- $d\mathbf{x} := \mathbf{x} - \mathbf{x}_0$ , in Komponentendarstellung  $d\mathbf{x} = (dx_1, dx_2, \dots, dx_n)^T$ .
- $dz := f'(\mathbf{x}_0)d\mathbf{x}$

Die durch

$$dz = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0) dx_j$$

beschriebene lineare Funktion mit den Variablen  $dx_1, \dots, dx_n$  heißt **totales Differential** (auch **vollständiges Differential**) von  $f$  in  $\mathbf{x}_0$ .

## Bemerkungen:

- Das totale Differential kann als Funktion  $df : \mathbb{R}^n \rightarrow \mathbb{R}$  aufgefasst werden mit

$$df(dx_1, \dots, dx_n) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0) dx_j.$$

- Vereinbart man  $\frac{\partial z}{\partial x_j} = \frac{\partial f}{\partial x_j}(\mathbf{x}_0)$ , so kann man das totale Differential auch schreiben:

$$dz = \sum_{j=1}^n \frac{\partial z}{\partial x_j} dx_j$$

1

## **Motivation:**

Wir wollen ein Analogon zur Differenz (in  $\mathbb{R}$ )

$$\Delta z := f(\mathbf{x}) - f(\mathbf{x}_0)$$

für Funktionen mehrerer Veränderlicher finden.



# I Oldies Dritte

## Idee:

Verwende die allgemeine Definition von  $f'(\mathbf{x})$ :

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + k(\mathbf{x}) \\ \Rightarrow f(\mathbf{x}) - f(\mathbf{x}_0) = \Delta z &= f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + k(\mathbf{x}). \end{aligned}$$

**Bemerkung:** Da  $|k(\mathbf{x})|$  sehr klein ist für  $|\mathbf{x} - \mathbf{x}_0|$  klein, ist  $\Delta z \approx f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$  dann eine gute Näherung.

**Definition:** (totales  
Verwende folgende

**Definition:** (totales Differential)

Verwende folgende Notation:

- $d\mathbf{x} := \mathbf{x} - \mathbf{x}_0$ , in Komponentendarstellung  $d\mathbf{x} = (dx_1, dx_2, \dots, dx_n)^\top$ .
- $dz := f'(\mathbf{x}_0)d\mathbf{x}$

Die durch

$$dz = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0) dx_j$$

beschriebene lineare Funktion mit den Variablen  $dx_1, \dots, dx_n$  heißt **totales Differential** (auch **vollständiges Differential**) von  $f$  in  $\mathbf{x}_0$ .

beschriebene lineare Funktion mit den Variablen  
totales Differential (auch vollständiges Differenzial)

## Bemerkungen:

- Das totale Differential kann als Funktion  $df : \mathbb{R}^n \rightarrow \mathbb{R}$  aufgefasst werden mit

$$df(dx_1, \dots, dx_n) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0) dx_j.$$

- Vereinbarung man  $\frac{\partial z}{\partial x_j} = \frac{\partial f}{\partial x_j}(\mathbf{x}_0)$ , so kann man das totale Differential auch schreiben:

$$dz = \sum_{j=1}^n \frac{\partial z}{\partial x_j} dx_j$$

1

# Taylor-Formel

## Motivation

- Sei eine  $f$  in  $a$  gegeben. Betrachte die Ableitung  $f'(a) = f'(a)$ , so mal, mal die Differentialquotient einer Funktion  $f: D \rightarrow \mathbb{R}$  ( $D \subseteq \mathbb{R}^n$ )
- Sei  $a$  ein Punkt der linearen Approximation  $f(a)$  von  $f(x)$  am Punkt  $a$ .  

$$f(x) \approx f(a) + \nabla f(a) \cdot (x - a)$$

$$f(x) = f(a) + \nabla f(a) \cdot (x - a) + \frac{1}{2} (x - a)^T H_f(a) (x - a) + \dots$$
- Folge:  $f$  ist eine  $p$ -te Ableitung an  $f'$  (so  $p$ -te Ableitung an  $f'$  sein)
- $f'(a) = 1$ , falls  $a$  eine  $p$ -te Ableitung an  $f'$  ist

$$f(x) = f(a) + \nabla f(a) \cdot (x - a) + \frac{1}{2!} \nabla^2 f(a) \cdot (x - a)^{\otimes 2} + \dots + \frac{1}{p!} \nabla^p f(a) \cdot (x - a)^{\otimes p} + \dots$$

## Vorbemerkung:

- Formalismus für Taylor-Polynom  $p$ -ten Grades recht aufwändig! Meist nur 1. und 2. Grad benötigt.
- Taylor-Polynom 1. Grades ist gleich Tangentenabbildung  $g(x)$ :

$$T_1(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0)$$

## Satz: (Taylor-Polynom 2. Grades)

Das Taylor-Polynom 2. Grades einer Funktion  $f$  an  $x_0$  kann in der Form

$$T_2(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2} (x - x_0)^T H_f(x_0) (x - x_0)$$

angegeben werden, wobei  $H_f$  die Hesse-Matrix von  $f$  ist. 2

## Satz (Zwischenwert)

Sei  $f: D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  eine  $p$ -mal diffbar und  $a, b \in D$  im Inneren von  $D$ .  
 Es gilt  $f(b) - f(a) = \int_a^b \nabla f(x) \cdot dx$ .

$$f(b) - f(a) = \int_a^b \nabla f(x) \cdot dx$$

## Exkurs: Abschätzung

$$|f(b) - f(a) - \int_a^b \nabla f(x) \cdot dx| \leq \int_a^b |\nabla f(x)| \cdot |dx|$$

## Beweisidee:

- Vielmehr: Taylor-Formel mit  $p=0$ :  

$$f(b) - f(a) = \int_a^b \nabla f(x) \cdot dx$$
- Abschätzung des Integrals (Cauchy-Schwarz'sche Ungleichung), Ersetze die erste Formel.
- Für die Abschätzung werde Mittelwertsatz für Vektorfelder auf  $F(x) = f(x) + \lambda(x - a)$  an.

## Wähle Vekturfeld

Es gilt  $f(b) - f(a) = \int_a^b \nabla f(x) \cdot dx$ . Sei  $\lambda$  ein Skalar, so dass  $F(x) = f(x) + \lambda(x - a)$  gilt.

$$F(b) - F(a) = \int_a^b \nabla F(x) \cdot dx = \int_a^b \nabla f(x) \cdot dx + \lambda(b - a) = f(b) - f(a) + \lambda(b - a)$$

## Satz: (Taylor-Formel im $\mathbb{R}^n$ )

Sei  $f: D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  eine  $(p+1)$ -mal stetig diffbare Funktion, sei weiter  $[a, a+h] \subseteq D$  ganz in  $D$  enthalten. Dann gilt die **Taylor-Formel**

$$f(a+h) = f(a) + \sum_{j=1}^p \frac{1}{j!} (h \cdot \nabla)^j f(a) + R(a, h)$$

Das Restglied  $R$  ist gegeben

$$R(a, h) = \int_a^{a+h} \frac{(1-s)^p}{p!} (h \cdot \nabla)^{p+1} f(a+sh) ds$$

Es gilt die Abschätzung

$$|R(a, h)| \leq \frac{|h|^{p+1}}{(p+1)!} \max_{s \in [0,1]} \sqrt{\sum_{i_1, \dots, i_{p+1}=1}^n |f_{i_1, \dots, i_{p+1}}(a+sh)|^2}$$

## Definition: (Taylor-Polynom $p$ -ten Grades im $\mathbb{R}^n$ )

Sei  $f: D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  eine  $(p+1)$ -mal stetig diffbare Funktion und  $[a, a+h] \subseteq D$  eine im Inneren von  $D$  liegende Strecke. Dann heißt

$$T_p(a, h) := f(a) + \sum_{j=1}^p \frac{1}{j!} (h \cdot \nabla)^j f(a)$$

**Taylor-Polynom  $p$ -ten Grades** der Funktion  $f(x)$  an der Stelle  $a$ .

## Bemerkung:

Setze  $x = a+h$  und  $x_0 = a$  (also  $h = x - x_0$ ), so lautet das Taylor-Polynom

$$T_p(x) = f(x_0) + \sum_{j=1}^p \frac{1}{j!} (x - x_0)^T \nabla^j f(x_0)$$

## Motivation:

- Setzt man in der allgemeinen Definition für Differenzierbarkeit von Abbildungen  $m = 1$ , so erhält man die Differenzierbarkeit einer Funktion  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$ .
- Nach der Idee von der linearen Approximation  $g(\mathbf{x})$  von  $f(\mathbf{x})$  kann man dann schreiben (sei  $\mathbf{x} = \mathbf{x}_0 + \mathbf{h}$ ):

$$g(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0) \cdot \mathbf{h} = f(\mathbf{x}_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0) h_j.$$

- **Frage:** Kann bei anderen Voraussetzungen an  $f$  bessere Approximation erreicht werden?
- Für  $n = 1$  hatten wir dafür die Taylor-Formel gefunden!

**Tools:**

- Der **Nabla-Operator** ist ein symbolischer Vektor-Operator:

$$\nabla := \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

- $\nabla$  angewandt auf eine Funktion  $f(x_1, \dots, x_n)$  ergibt (wie zuvor verwendet)

$$\nabla f(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = \text{grad} f.$$

- Formale Skalarmultiplikation mit Vektor  $\mathbf{h} = (h_1, \dots, h_n)^\top \in \mathbb{R}^n$  ergibt

$$\mathbf{h} \cdot \nabla = h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n} = \sum_{j=1}^n h_j \frac{\partial}{\partial x_j}.$$

- Anwendung auf  $f$ :

$$(\mathbf{h} \cdot \nabla) f(\mathbf{x}) = \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}.$$

- Die  $k$ -te Potenz des Operators sei zu verstehen:

$$(\mathbf{h} \cdot \nabla)^k = \sum_{i_1, i_2, \dots, i_k=1}^n h_{i_1} h_{i_2} \dots h_{i_k} \frac{\partial^k}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}.$$

- Entsprechend angewendet auf  $f$ :

$$(\mathbf{h} \cdot \nabla)^k f(\mathbf{x}) = \sum_{i_1, i_2, \dots, i_k=1}^n h_{i_1} h_{i_2} \dots h_{i_k} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}.$$

$$(\mathbf{h} \cdot \nabla)^k f(\mathbf{x}) = \sum_{i_1, i_2, \dots, i_k=1}^n h_{i_1} h_{i_2} \cdots h_{i_k} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}}.$$

### Weitere Vereinfachung:

Falls alle partielle Ableitungen stetig sind, kann der Satz von Schwarz angewendet werden:

$$(\mathbf{h} \cdot \nabla)^k f(\mathbf{x}) = \sum_{i_1 + i_2 + \cdots + i_n = k} h_1^{i_1} h_2^{i_2} \cdots h_n^{i_n} \left( \frac{\partial}{\partial x_1} \right)^{i_1} \left( \frac{\partial}{\partial x_2} \right)^{i_2} \cdots \left( \frac{\partial}{\partial x_n} \right)^{i_n} f(\mathbf{x}).$$

**Satz:** (Taylor-Formel im  $\mathbb{R}^n$ )

Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  eine  $(p + 1)$ -mal stetig diff'bare Funktion; sei weiter  $[\mathbf{a}, \mathbf{a} + \mathbf{h}] \subset D$  ganz in  $D$  enthalten. Dann gilt die **Taylor-Formel**

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \sum_{j=1}^p \frac{1}{j!} (\mathbf{h} \cdot \nabla)^j f(\mathbf{a}) + R(\mathbf{a}, \mathbf{h}).$$

Das Restglied  $R$  ist gegeben

$$R(\mathbf{a}, \mathbf{h}) = \int_0^1 \frac{(1-s)^p}{p!} (\mathbf{h} \cdot \nabla)^{p+1} f(\mathbf{a} + s\mathbf{h}) ds.$$

Es gilt die Abschätzung

$$|R(\mathbf{a}, \mathbf{h})| \leq \frac{|\mathbf{h}|^{p+1}}{(p+1)!} \max_{0 \leq s \leq 1} \sqrt{\sum_{i_1, i_2, \dots, i_{p+1}=1}^n |f_{x_{i_1} x_{i_2} \dots x_{i_{p+1}}}(\mathbf{a} + s\mathbf{h})|^2}.$$



**Definition:** (Taylor-Polynom  $p$ -ten Grades im  $\mathbb{R}^n$ )

Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  eine  $(p+1)$ -mal stetig diff'bare Funktion und  $[\mathbf{a}, \mathbf{a}+\mathbf{h}] \subset D$  eine im Inneren von  $D$  liegende Strecke. Dann heißt

$$T_p(\mathbf{a} + \mathbf{h}) := f(\mathbf{a}) + \sum_{j=1}^p \frac{1}{j!} (\mathbf{h} \cdot \nabla)^j f(\mathbf{a})$$

**Taylor-Polynom**  $p$ -ten Grades der Funktion  $f(\mathbf{x})$  an der Stelle  $\mathbf{a}$ .

**Bemerkung:**

Setze  $\mathbf{x} = \mathbf{a} + \mathbf{h}$  und  $\mathbf{x}_0 = \mathbf{a}$  (also  $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$ ), so lautet das Taylor-Polynom

$$T_p(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{j=0}^p \frac{1}{j!} (\mathbf{h} \cdot \nabla)^j f(\mathbf{x}_0).$$

**Satz:** (Mittelwertsatz)

Ist  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  einmal stetig diff'bar und  $[\mathbf{a}, \mathbf{a} + \mathbf{h}]$  eine im Inneren von  $D$  liegende Strecke. Dann gibt es  $0 < \theta < 1$ , so dass gilt

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \sum_{j=1}^n h_j f_{x_j}(\mathbf{a} + \theta \mathbf{h}).$$

Es gilt die Abschätzung

$$|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a})| \leq |\mathbf{h}| \max_{0 \leq s \leq 1} \sqrt{\sum_{j=1}^n |f_{x_j}(\mathbf{a} + s\mathbf{h})|^2}.$$

**Beweisidee:**

- Verwende Taylor-Formel mit  $p = 0$ :

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \int_0^1 (\mathbf{h} \cdot \nabla) f(\mathbf{a} + s\mathbf{h}) ds.$$

- Abschätzung des Integrals (Cauchy-Schwarzsche Ungleichung) liefert die erste Formel.
- Für die Abschätzung wende Mittelwertsatz für eine Veränderliche auf  $F(s) := f(\mathbf{a} + s\mathbf{h})$  an.

### Vorbemerkung:

- Formalismus für Taylor-Polynom  $p$ -ten Grades recht aufwändig!  
Meist nur 1. und 2. Grad benötigt.
- Taylor-Polynom 1. Grades ist gleich Tangentenabbildung  $g(\mathbf{x})$ :

$$T_1(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0).$$

### Satz: (Taylor-Polynom 2. Grades)

Das Taylor-Polynom 2. Grades einer Funktion  $f$  an  $\mathbf{x}_0$  kann in der Form

$$T_2(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^\top H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

angegeben werden, wobei  $H_f$  die Hesse-Matrix von  $f$  ist.

## Erinnerung: Ableitungsmatrix und Lineare Approximation

**Definition 1** Sei  $f: D \rightarrow \mathbb{R}^m$  eine Funktion, die in einem Punkt  $a \in D$  diffbar ist. Dann ist die Ableitungsmatrix von  $f$  in  $a$  definiert durch

$$Df(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

Die lineare Approximation von  $f$  in  $a$  ist die Funktion

$$L_f(a, x) = f(a) + Df(a)(x - a)$$

Die Taylor-Entwicklung von  $f$  in  $a$  ist die Funktion

$$T_f(a, x) = f(a) + Df(a)(x - a) + \frac{1}{2} D^2 f(a)(x - a, x - a) + \dots$$

## Totales Differential

**Satz 1** Sei  $f: D \rightarrow \mathbb{R}^m$  eine Funktion, die in einem Punkt  $a \in D$  diffbar ist. Dann ist das totale Differential von  $f$  in  $a$  definiert durch

$$df_a = Df(a)(x - a)$$

**Definition 2** Sei  $f: D \rightarrow \mathbb{R}^m$  eine Funktion, die in einem Punkt  $a \in D$  diffbar ist. Dann ist das totale Differential von  $f$  in  $a$  definiert durch

$$df_a = Df(a)(x - a)$$

**Beispiel** Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  die Funktion  $f(x, y) = x^2 + y^2$ . Dann ist das totale Differential von  $f$  in  $a = (1, 1)$  definiert durch

$$df_a = 2x dx + 2y dy = 2 dx + 2 dy$$

## Taylor-Formel

**Satz 2** Sei  $f: D \rightarrow \mathbb{R}^m$  eine Funktion, die in einem Punkt  $a \in D$   $k$ -mal diffbar ist. Dann ist die Taylor-Formel von  $f$  in  $a$  definiert durch

$$T_f(a, x) = f(a) + Df(a)(x - a) + \frac{1}{2} D^2 f(a)(x - a, x - a) + \dots + \frac{1}{k!} D^k f(a)(x - a, \dots, x - a)$$

**Definition 3** Sei  $f: D \rightarrow \mathbb{R}^m$  eine Funktion, die in einem Punkt  $a \in D$   $k$ -mal diffbar ist. Dann ist die Taylor-Formel von  $f$  in  $a$  definiert durch

$$T_f(a, x) = f(a) + Df(a)(x - a) + \frac{1}{2} D^2 f(a)(x - a, x - a) + \dots + \frac{1}{k!} D^k f(a)(x - a, \dots, x - a)$$

**Beispiel** Sei  $f: \mathbb{R} \rightarrow \mathbb{R}$  die Funktion  $f(x) = e^x$ . Dann ist die Taylor-Formel von  $f$  in  $a = 0$  definiert durch

$$T_f(0, x) = 1 + x + \frac{1}{2} x^2 + \dots + \frac{1}{k!} x^k$$